大学生学业内卷的测量及其影响*

刘莹莹　谢昕彤　张顺民

（浙江大学心理与行为科学系　杭州　310058）

摘　要 学业内卷虽有可能危害高校人才培养，但尚缺乏可靠的测量工具进行检验。本文通过3个研究开发了大学生学业内卷问卷并验证其有效性。研究1结合文献分析、生活观察和访谈构建了初始题库，并基于338名大学生的试测结果筛选了题目。研究2基于大样本（*N*＝3000）的探索性因子分析和独立样本（*N*＝571）的验证性因子分析确立了16个条目，三因子（舍心逐物、竞争失格和脱实务虚）的最终版问卷。3000人的样本显示，多数学生表现出较高的学业内卷。高学业内卷个体的零和信念更强、特质焦虑更高、生活满意度更低、睡眠质量更差，但创造力倾向却没有更高。研究3基于99人的数据发现问卷重测信度为0.83。本研究不仅开发了可靠的测量工具，还发现大学生学业内卷普遍存在且危害明显。

关键词 学业内卷，量表开发，焦虑，睡眠，创造力

The Measurement and Influence of Colleges' Academic Involution

Liu Yingying　Xie Xintong　Zhang Shunmin

(Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310058)

Abstract: Academic involution may harm the cultivation and development of college students, but there has not been a reliable measurement tool to assess it. This paper developed a Colleges' Academic Involution Scale (CAIS) and examined its reliability and validity with 3 studies. Study 1 generated a 31-item pool based on literature review, daily cases, and interview, and filtered items based on a 338-undergraduate sample. Study 2 confirmed a 16-item final version CAIS, which consisted of three dimensions: unwilling hardworking, excessive competition, and surface learning, based on a large sample (*N* = 3000) and an independent sample (*N* = 571). Based on the 3000-undergraduate sample, more than 60% of college students are involved in academic involution. Specifically, individuals with high scores in the CAIS showed stronger zero-sum belief, higher trait anxiety, lower life satisfaction, and poorer sleep quality, but not greater creative potential. Study 3 revealed that the test-retest reliability of the final version scale reached 0.83 based on a new sample

---

* 国家自然科学基金青年项目(32100870)
通讯作者：张顺民，E-mail：shunmin@zju.edu.cn

($N = 99$). The CAIS could be a reliable and effective tool for future research exploring harms, causes, and ways to mitigate academic involution.

Keywords academic involution, scale development, trait anxiety, sleep, creativity

1 引言

内卷，作为人们对“过度竞争”和“努力贬值”的一种描述，在近年受到了社会广泛而持续的关注。2020 年流传的一张照片引发人们对内卷的热议。照片中，一名大学生边骑自行车、边在笔记本电脑上学习。这一画面让人们惊惑：高校的学业竞争已经激烈到如此程度了吗？随着内卷一词的流行，越来越多深入的疑问接踵而至。学业内卷现象真的存在吗？高校学生的内卷程度如何？内卷对于人才的培养有何影响？造成学业内卷现状的原因又是什么？要深入调查以上诸多问题，明确的内卷定义和有效的测量工具不可或缺。

内卷（involution）一词由人类学家 Goldenweiser 于 1936 年引入至社会科学领域，用于描述“文化模式在临近某种确定形态时，既没有办法稳定下来，也没有办法使自己转变到新的形态，取而代之的是不断地在内部变得更加复杂”[2]的现象。Geertz(1963)则援引“内卷”描述爪哇地区的农业现象：无数的劳动力集中在有限的水稻生产中，尽管单位面积的产量有增加，但人们的生活水平却只有非常小的提高。黄宗智(2002)是首位将“内卷”概念引入国内研究的学者，其在专著中对内卷做了注解，认为内卷出现于劳动力的边际产量开始递减之后。尽管这种界定与 Geertz 的定义略有差异(刘世定, 邱泽奇, 2004)，但黄宗智所想表达的内容与 Geertz 仍有共通之处。黄宗智(2021)所强调的内卷特点有：(1) 越来越高的劳动投入、(2) 越来越低的边际回报、(3) 形成顽固的封闭体系、(4) 抵御甚至排除质变。虽然多数学者认为内卷有着消极的影响，少数学者仍尝试过对内卷做出积极的解读(郭继强, 2007)，认为 Geertz 曾将内卷描述为一个“自我战胜的过程”。然而，这种积极解读是对 Geertz 内卷概念的曲解。“自我战胜的过程”对应的原文是“self-defeating process”[3]。Self-defeating 意为“弄巧成拙”或“自我挫败”，而非“自我战胜”。在 Geertz 的描述中，内卷中的爪哇农业生产者是无奈的、被迫的。他们在有限的土地中不断投入更多的劳动力，却始终只能勉强维持生存、看不到改变现状的希望(McCullough, 2019)。

综上可知，内卷的基本特点有：(1) 既定框架的限制、(2) 不断攀升的投入、(3) 缺乏

[2] Culture patterns that in reaching a definitive form, do not evolve into new patterns, but continue to develop only in the direction of internal complexities (Geertz, 1963, P81).
[3] It is this ultimately self-defeating process that I have proposed to call “agricultural involution.” (Geertz, 1963, P80).

实质收益或发展。Goldenweiser 的“文化模式的某种确定形态”、Geertz 的“有限的水稻生产”和黄宗智(2021)所说的“封闭体系”都构成了某种既定的框架。而 Goldenweiser 的“在内部变得更加复杂”、Geertz 的“无数劳动力集中”以及黄宗智(2021)所说的“越来越高的劳动投入”都体现了投入的不断攀升。从个体的角度而言，内卷的体现是有限的实质收益；而从系统的角度而言，内卷则体现的是缺少实质的发展。Geertz 的描述更体现个体的角度，即“人们的生活水平却只有非常小的提高”。而 Goldenweiser 所描述的现象更体现系统的角度，即“文化模式的某种确定形态”无法“转变到新的形态”。相比之下，黄宗智(2002)的描述则对个体视角和系统视角都有所体现。从个体视角而言，他虽然承认内卷现象下个体仍可以通过增加劳动时间来提高收入，但由于边际产量递减，这种收益增加终究是有限的。从系统视角而言，他也强调这种收入提高应区别于“发展”，因为整个系统并没有出现劳动生产率的提升。综合以上分析，内卷的定义可提炼为：一种在既定框架内部投入不断攀升却缺少实质收益或发展的现象。

教育领域是否存在内卷现象？答案或许是肯定的。名校的深造机会、知名企业的就业岗位，这些优质资源的名额限制便构成了一个封闭的空间，而名额争夺实则是这种封闭空间内的零和博弈(陈诚, 包雷, 2022)。为了不在零和竞争中落败，每个学生都不得不追加更多的学习投入。然而，这些投入却并不一定能给个人带来实质收益。一项基于中国教育追踪调查的研究发现，随着中学生课外补习时间的整体提升，个体补习时间对成绩提升效果逐渐被削弱，而焦虑和抑郁却与日俱增(张骞, 高雅仪, 2022)。从学习或成长的角度而言，过长的学习时间会降低学习效率(庞守兴, 李书肖, 2010)，而学业焦虑也容易引起学习倦怠(高明, 2014)。此外，张骞和高雅仪还发现课外补习时间的增加会降低学生的睡眠质量，而睡眠不足则有害于大脑健康、不利于学生的成长(Eide et al., 2021)。另一方面，由于学校难以全面、准确地测量能力，而只能以成绩概括，部分学生便只求“成绩”不求“成长”。为了提高绩点，一些学生会选择专业关联性较低却相对容易的课程(赵婷婷, 胡艳婷, 2022)。在许多高校，学生之间甚至还流传着各种帮助“刷绩点”的选课指南。在这种情况下，即使学生投入大量的学习时间也不一定学到了实质的知识。从人才培养的系统角度而言，当下“唯学历”、“唯名校”的既定评价体系也缺乏实质发展，不利于国家保持合理的人才培养结构(熊丙奇, 2021)。由此可见，人们对教育内卷的抱怨未必是空穴来风。

教育内卷可能有着诸多的危害。学者们的担忧之一是，高校学生可能在内卷中趋于同质，失去个性和创造力(夏骏, 2012)。许多大学生仍然延续着应试的心态、追求分数和升学率(陈雪飞, 2019)，而内卷只会加重这种心态。高层次人才的培养也可能受到内卷的影响。在博士

生论文发表制度“内卷化”的情况下，学术论文可能出现“有数量增长而无实质创新”的现象(赵祥辉, 2021)。这显然不符合全国研究生教育会议中“培养高层次的拔尖创新人才”的要求。内卷还可能对学生的心理健康、家庭的教育投入带来负担。严重的内卷不仅容易使年轻一代丧失学习动机，还可能导致焦虑、抑郁甚至自杀和暴力现象(戴子涵, 2022)。内卷也可能波及整个家庭，造成“家长教育投入增加、焦虑却有增无减”的困境(陈诚, 包雷, 2022)。在更宏观的意义上，教育内卷不仅导致社会资源的浪费，也不利于国家的长期发展。

虽然对教育内卷的深入研究势在必行，但可靠的内卷测量工具却仍不可得。目前虽已有三个研究尝试编制了学业内卷问卷(Yan et al., 2022; Yang et al., 2023; Yi et al., 2022)，但它们都存在着内卷定义不明、内容效度欠佳或信效度证据缺失等诸多问题。为了提供可靠的测量工具，本研究从提炼内卷定义出发，进行问卷编制并系统地检验了问卷的信效度。学业内卷的操作性定义为，个体在应试或排名竞争中投入攀升却缺乏知识、能力或人格成长的现象。根据操作性定义可知，个体的学业内卷可能有外部和内部两种相互关联的具体表现。对外，内卷倾向较高的学生会为了高绩点或高排名而过度投入或过度竞争。对内，高内卷的学生则缺乏能力或人格的成长。研究 1 将根据以上表现，结合内卷定义、文献总结、生活观察和学生访谈等方式构建有代表性的初始问卷。研究 2 将利用大样本形成最终的内卷问卷并利用零和信念、特质焦虑、生活满意度、创造力倾向及睡眠质量作为关联校标进行效度检验(王琛, 2013)。本研究预期学业内卷应该与零和信念、特质焦虑呈正相关，与生活满意度和睡眠质量呈负相关。由于内卷意味着缺乏实质发展和收益，因此高内卷的学生不会呈现出更高的创造力倾向。研究 3 将在新的样本中检验最终版问卷的重测信度。

2 研究 1：问卷题目的编制与筛选

2.1 研究目的

研究 1 结合内卷定义、生活观察和访谈等方式构建大学生学业内卷问卷的初始题库，并对初始问卷开展定量分析，删除存在严重问题的题目，形成初始版大学生学业内卷问卷。

2.2 研究方法

2.2.1 题目编制与问卷结构

根据其操作性定义推测，学业内卷可能存在外部和内部两种相互关联的表现。在外部表现为以获得高绩点或高排名为目标的过度投入或过度竞争。在内部为缺乏能力或人格的成长。研究 1 依据文献分析所提炼的学业内卷定义，结合生活中的观察与访谈，筛选现实中有如上体现的实例，构建了问卷的初始题库。例如，体现“过度投入”的题目可能包括“只要

能让绩点排名靠前，我愿意在某一门课程上投入大量时间”。关注“过度竞争”的题目可能包括“我会担心休息时被身边的同龄人超越”。反映“缺乏能力成长”可能包括“我参加很多收获不大的社会实践或实习，主要是为了让我的履历更丰富一些”。侧重“缺乏人格成长”的题目可能包括“我没有真正喜欢的工作，只是努力不在各种竞争中落后”。此外，在初始题库中设计了少量反向计分题。例如，“过度投入”的反向计分题：“我作息规律，不会因为考试或者排名而熬夜学习”。“过度竞争”的反向计分题：“即使不用与人竞争，我也会保持良好的学习状态”。“缺乏能力成长”的反向计分题：“我努力让自己写的论文都是有价值的，而不仅是想方设法拿到高分”。“缺乏人格成长”的反向计分题：“我会根据自己的兴趣选择课程，而不是看老师给分的高低”。

初始题库共 31 题，其中 8 道反向计分题。均为 5 点李克特量表评分（1：非常不符合~5：非常符合），评分越高代表被试认为自己与该项目描述的情况越符合。

2.2.2 研究对象

研究 1 以在读大学生为对象进行施测。共计在线发放问卷 373 份，其中 35 名被试未按照要求完成问卷中添加的注意力检测题，其数据未被纳入分析。保留有效问卷 338 份，问卷有效回收率为 90.62%。被试平均年龄为 20.52±1.53 岁，其中男性 117 名(34.62%)，包含了大一(4.44%)、大二(29.59%)、大三(39.05%)、大四(15.98%)和其他(10.95%)年级的学生。

2.3 研究结果

2.3.1 项目分析

为了检验初始题库中各题目的区分度，根据被试的问卷总分排序（计算总分之前对反向计分题得分进行了转换），按照 27%分位数分为高低两组，对每道题进行独立样本 T 检验，发现高分组(N = 91)与低分组(N = 88)在 28 道题上差异显著($p_{\text{Bonferroni correction}} < 0.01$)。另外 3 道题差异不显著（题目 7、12、22），表示其区分度较低。

为了检验初始题库中单个题目与其他题目之间的异质性，计算单个题目与其他所有题目的总分之间相关性(corrected item-total correlations)。基于以往文献，该相关性指标在 0.30~0.70 之间是可接受的(Vagos et al., 2016)。根据这一标准，初始题库中有 6 道题低于 0.30（题目 1、2、7、12、17、22，相关大小分别是：$r_1 = 0.24$, $r_2 = 0.28$, $r_7 = -0.15$, $r_{12} = 0.09$, $r_{17} = 0.23$, $r_{22} = -0.08$），表示其异质性较高。

为了避免存在严重问题的题目影响后续探索性因子分析，在进一步的数据分析之前，删除区分度不高且异质性较高的题目（题目 7、12、22，均为反向计分题）。

2.3.2 探索性因子分析

对保留的 28 道题进行探索性因子分析，初步探测问卷的因子结构。因子分析可行性结果显示：KMO 值为 0.88，高于经验标准 0.80，表示变量间的共同因素较多。巴特利特球形检验 χ^2= 3174.84 (df = 378, p < 0.001)，表明数据适合进行探索性因子分析。

使用主成分分析法提取因子，根据特征值大于 1 的标准提取因子，并通过 Promax 最优斜交法旋转因子。根据特征值大于 1 的标准，建议提取 6 个因子，总方差解释率为 57.14%。结合以往文献(Chen et al., 2006; Olatunji et al., 2007)与研究 1 的目的（删除问题严重的题目，同时尽可能多的保留性能可接受的题目），确定以下保留因子的标准：（1）题目在该因子上的载荷大于 0.3，（2）该因子至少包含 3 道题。结果表明，各题目的因子载荷均超过 0.3，符合标准（1）；但因子 6 仅包含 1 道题（第 14 题），不符合上述因子保留标准（2），考虑删除。另外，初始题库中为不同维度设计的 5 道反向计分题（题目 1、17、18、19、21）单独构成了一个因子，失去了作为不同维度反向计分题的作用，考虑删除。

2.4 大学生学业内卷问卷初始版

基于上述项目分析和探索性因子分析结果，初步考虑删除 1、7、12、14、17、18、19、21、22 题。但是结合先前文献、生活观察与访谈结果，第 1 题："我会根据自己的兴趣选择课程，而不是看老师给分的高低"，在讨论内卷时经常被提及。因此，在初始题库中保留了这一题目，并将其修改为正向计分题"选课时，我优先选择容易拿高分的课程，以便取得更高的绩点"，进一步探究该题表现。同时，对其他保留的题目在维持原意不变的情况下，对其内容稍作修改，使其表达更简洁直接。至此，研究 1 形成了包含 23 道题的初始版大学生学业内卷问卷。

3 研究 2：问卷修订及信效度检验

3.1 研究目的

研究 2 利用大样本对初始版问卷进行修订，并在新样本中检验修订版问卷的信效度，以确定最终版大学生学业内卷问卷。

3.2 研究方法

3.2.1 研究对象

研究 2 收集了两个样本数据。样本 1 用于探索性因子分析对初始版问卷进行修订，而样本 2 用于验证性因子分析检验修订版问卷因子维度的稳定性。以上两个步骤共同确定大学生学业内卷问卷的最终版。基于最终版问卷，结合样本 1 收集的校标数据，研究 2 进一步探索了学业内卷与零和信念、特质焦虑、生活满意度、创造力倾向和睡眠情况的关系，进一步

确定问卷的结构效度。

样本 1：委托见数平台(https://www.credamo.com/#/)收集 3000 份有效问卷，所有被试均按照要求完成了问卷中设置的注意力检测题。样本涵盖了全国 32 个省、市、自治区。被试平均年龄为 20.80±1.58 岁，其中男性 919 名(30.63%)，包含了 985(11.33%)、211(19.07%)、一本（非 985 和 211，29.10%）、二本(24.90%)、三本(4.73%)、专科(9.90%)等各类院校的学生，由大一(12.33%)、大二(23.30%)、大三(29.00%)、大四(24.70%)和其他年级(10.67%)的学生组成。

样本 2：自主招募大学生进行施测，共计发放 580 份问卷，有 9 名被试未通过注意力检测题，回收有效问卷 571 份，问卷回收率为 98.44%。样本涵盖了全国 30 个省、市、自治区，被试平均年龄为 21.19±1.73 岁，其中男性 166 名(29.07%)，包含了 985(11.91%)、211(15.41%)、一本（非 985 和 211，33.10%）、二本(27.50%)、三本(5.25%)、专科(6.30%)等各类院校的学生，由大一(10.86%)、大二(23.47%)、大三(31.87%)、大四(22.94%)和其他年级(10.86%)的学生组成。

3.2.2 研究工具

（1）大学生学业内卷问卷：样本 1 完成大学生学业内卷问卷初始版（23 道题目）。样本 2 完成大学生学业内卷问卷修订版，修订版由初始版问卷经探索性因子分析后得到的 17 道题目组成。大学生学业内卷问卷的初始版和修订版均采用李克特 5 点计分，从“非常不符合”到“非常符合”分别计为“1”到“5”分，评分越高代表被试的学业内卷程度越高。

（2）零和信念量表：采用 Crocker 和 Canevello (2008)编制的 6 个题目组成的量表（如：一个人的成功取决于另一个人的失败），该量表采用李克特 7 点计分，评分越高表示被试的零和信念越强。在样本 1 中零和信念量表的 α 一致性系数为 0.72。

（3）特质焦虑量表：该量表包含 20 个题目（如：我感到神经过敏和不安），采用 1-4 级评分（1-完全没有，2-有些，3-中等程度，4-非常明显），被试得分越高表示其特质焦虑水平越高(汪向东 等, 1999)。在样本 1 中该量表的 α 一致性系数为 0.88。

（4）生活满意度量表：该量表包含 5 个题目（如：我满意自己的生活），采用李克特 7 点计分，评分越高表示被试的整体生活满意度越高(Diener et al., 1985)。在样本 1 中该量表的 α 一致性系数为 0.87。

（5）威廉姆斯创造力倾向量表：该量表包含好奇心、冒险性、想象力和挑战性四个维度(Liu et al., 2018)，共 50 个题目（如：我喜欢仔细观察我没有见过的东西，以了解详细的情形），其中 8 个题目为反向计分题，在计算总分之前转换反向计分题得分，总分越高表示

被试的创造力潜能越高。在样本 1 中该量表的 α 一致性系数为 0.87。

（6）睡眠：采用自编的 2 个题目测量被试的睡眠时间（您平均每天睡多少个小时？您觉得您的睡眠时间充足吗？），1 个题目测量被试是否存在失眠问题（您有失眠问题吗？）。

（7）家庭社会经济地位：采用 10 级量表的形式麦克阿瑟梯子测量被试主观的家庭社会经济地位(Goodman et al., 2001)，该信息仅作为基本信息，并未纳入后续的数据分析。

3.2.3 共同方法偏差检验

采用探索性因子分析对可能存在的共同方法偏差进行检验。整合样本 1 被试完成的大学生学业内卷问卷初始版、零和信念量表、特质焦虑量表、生活满意度量表、威廉姆斯创造力倾向量表和睡眠情况调查问卷中所有项目进行探索性因子分析，析出的因子数目大于 1 且第一个公因子解释率为 12.11%，小于 40%，表明样本 1 的数据不存在严重的共同方法偏差。

3.3 研究结果

3.3.1 项目分析

为了检验初始版问卷各题目的区分度，根据样本 1 被试在初始版学业内卷问卷上的总分进行排序，按照 27%分位数分为高低两组，对每道题进行独立样本 *T* 检验，发现高分组(N = 803)与低分组(N = 842)在所有题目上均差异显著($p_{\text{Bonferroni correction}} < 0.01$)，表示各题目均具有良好的区分度。

为了检验初始版问卷中单个题目与其他题目之间的异质性，计算单个题目与其他所有题目的总分之间相关性(corrected item-total correlations)。采用与研究 1 相同的标准，初始版问卷有 3 道题的相关性低于 0.30（题目 2、13、29，相关大小分别是：$r_2 = 0.17$，$r_{13} = 0.28$，$r_{29} = 0.19$），表示这 3 道题目的异质性较高。

为了确保学业内卷问卷具有较高的信效度，如果题目质量在研究 2 中仍未达到可接受的标准，表示该题不能较好反映学生的学业内卷程度，考虑删除。因此，在进行后续分析之前，删除异质性较高的 3 道题（题目 2、13、29）。

3.3.2 探索性因子分析

对初始版问卷保留的 20 道题进行 KMO 和巴特利特球形检验。结果显示 KMO 值为 0.91，高于经验标准 0.80，表示变量间的共同因素较多。巴特利特球形检验 $\chi^2 = 14507.78$ ($df = 190$, $p < 0.001$)，表明适合进行因子分析。

使用主成分分析法提取因子，根据因子特征值大于 1 和碎石图标准共同决定因子提取数目，通过 Promax 最优斜交法旋转因子。特征值大于 1 和碎石图的结果均建议提取三个因子（各题目的因子载荷见表 1）。三因子模型解释了总方差的 43.69%。结合以往文献(Chen et

al., 2006)与研究 2 的目的（确保问卷的信效度，保留性能较好的题目），确定以下保留因子的标准：（1）题目在该因子上的载荷大于 0.4，（2）该因子至少包含 3 道题。结果表明，各因子下至少包含 3 道题，符合因子保留标准（2）；但其中 3 道题（题目 3、5、26）在任一因子上的载荷均小于 0.4，不符合上述因子保留标准（1），考虑删除。

保留的 17 道题分布在 3 个因子中。因子 1 包含 7 道题，将其命名为“舍心逐物”——努力追求外在结果而丧失学习初心。因子 2 包含 5 道题，将其命名为“竞争失格”——因资源有限而过度竞争。因子 3 包含 5 道题，将其命名为“脱实务虚”——不研究如何求学问而一味追求高绩点。以此作为大学生学业内卷问卷修订版，在样本 2 中检验修订版问卷因子结构的稳定性。

表 1　大学生学业内卷问卷的因子载荷矩阵

题目	舍心逐物	竞争失格	脱实务虚
4. 除了应对考试，我不知道自己所学的知识有什么用。	0.81		
6. 我偏好做题而非研读或思考理论，因为做题更有利于提高分数。	0.51		
15. 我虽然看起来在努力，但好像没有学到什么。	0.57		
23. 我努力学习，但不是因为我真的热爱学习本身。	0.74		
25. 我没有真正喜欢做的工作，只是努力不在各种竞争中落后。	0.72		
27. 当确定跨过了“胜过竞争对手”的门槛后，我便不会再进行任何额外的学习。	0.57		
31. 我们越来越努力了，但生活似乎并没有因此变好。	0.58		
16. 我会担心自己休息时被身边的同龄人超越。		0.76	
20. 我忙着提升竞争力，没有时间做自己真正想做的事。		0.41	
24. 如果发现有同学在学习而自己却在休息，我会非常地自责。[a]		0.72	
28. 竞争中暂时的领先或落后经常影响我的心情和学习状态。		0.56	
30. 我学得很累，但很少有时间休息。		0.68	
1. 选课时，我优先选择容易拿高分的课程，以便取得更高的绩点。			0.60
8. 为了加分（课程、保研等），我会参加实质意义不大、徒具形式的活动。			0.68
9. 我参加很多收获不大的实践或实习，主要是为了让履历更丰富一些。			0.67
10. 完成课程论文时，我会故意超出规定的字数，以期得到更好的成绩。			0.51
11. 只要能让绩点排名靠前，我愿意在某一门课程上投入大量的时间。			0.68

注：因子载荷小于 0.4 都未显示

[a] 该题目在样本 2 中进行验证性因子分析时删除。

3.3.3 验证性因子分析

为检验样本 1 中得到的修订版问卷因子结构是否稳定，使用样本 2 进行验证性因子分析。验证性因子分析的初步结果显示第 16 题与第 24 题的残差项之间有较强相关(modification index = 41.97)。从题目内容上看，两个题目测量内容重叠程度较高，均侧重于测量被试是否

因过度竞争而影响休息（第 16 题：我会担心自己休息时被身边的同龄人超越。第 24 题：如果发现有同学在学习而自己却在休息，我会非常地自责。）。从相关程度上看，第 16 题与第 24 题存在较强的正相关($r = 0.54$, $p < 0.001$, 95% CI = [0.48, 0.60])。因此，删除因子载荷相对较小的题目（第 24 题）。

进一步对保留的 16 道题进行验证性因子分析，模型拟合指标为：$\chi^2 = 353.45$, $df = 101$, GFI = 0.93, RMSEA = 0.07, CFI = 0.90。各项指标均在可接受的范围内，表明 16 道题、三因子结构组成的大学生学业内卷问卷是稳定的，并将此确定为最终版大学生学业内卷问卷。

3.3.4 内部一致性分析

基于样本 2 的数据对最终版问卷内部结构的一致性进行了分析。结果表明，三个子维度均与总分显著正相关($r_{舍心逐物} = 0.85$, $p < 0.001$, 95% CI = [0.82, 0.87]; $r_{竞争失格} = 0.80$, $p < 0.001$, 95% CI = [0.77, 0.83]; $r_{脱实务虚} = 0.73$, $p < 0.001$, 95% CI = [0.69, 0.77])，三个子维度之间强相关($r_{舍心逐物-竞争失格} = 0.49$, $p < 0.001$, 95% CI = [0.43, 0.55]; $r_{舍心逐物-脱实务虚} = 0.34$, $p < 0.001$, 95% CI = [0.27, 0.41]; $r_{竞争失格-脱实务虚} = 0.54$, $p < 0.001$, 95% CI = [0.48, 0.60])。进一步分析了最终版问卷题目的内部一致性。结果显示，总问卷的 α 一致性系数是 0.86，舍心逐物、竞争失格与脱实务虚三个子维度的 α 一致性系数分别是 0.83、0.72 和 0.70，均达到可接受的水平。上述结果表明最终版问卷内部结构一致、题目信度良好。

3.3.5 结构效度

基于最终版问卷结构，计算了样本 1 被试的内卷总分、三个子维度得分，并将其分别与零和信念、焦虑、生活满意度、创造力倾向、睡眠情况进行相关分析。结果显示（见表 2），内卷总分与各校标的相关关系与预期一致，高内卷者的零和信念更强、焦虑水平更高、生活满意度更低、睡眠时长更短、睡眠更不充足、失眠问题更严重，而创造力倾向却没有更高。进一步利用贝叶斯统计推断是否可以接受学业内卷与创造力没有相关的零假设。结果显示，贝叶斯因子 $BF_{01} = 10.40$，表明在零假设下出现当前数据的可能性是在备择假设下的 10.40 倍。根据分类标准(Jeffreys, 1961)，有中等强度的证据支持接受零假设。三个子维度与多数校标的相关也表现出与总分类似的模式，但三者与创造力倾向的关系却表现出了明显的分化。高舍心逐物的个体有着更低的创造力倾向，而高竞争失格、高脱实务虚的个体却都有着更高的创造力倾向。此外，相对于舍心逐物和竞争失格，脱实务虚与焦虑、睡眠时长和失眠问题的相关明显更弱，与睡眠充足、生活满意度的相关接近于 0 (睡眠充足：$BF_{01} = 12.25$，有较强的证据支持接受零假设；生活满意度：$BF_{01} = 34.79$，有较强的证据支持接受零假设)。总体而言，学业内卷问卷及三个子维度与各校标变量的关系，符合高内卷者过度竞争、关注

成绩或排名而忽略自身实质成长的行为模式，表明最终版大学生学业内卷问卷具有良好的结构效度。

表 2　大学生学业内卷问卷（最终版）的结构效度

变量	*M (SD)*	内卷总分	舍心逐物	竞争失格	脱实务虚
1.零和信念	23.21(6.46)	0.38^{***} [0.35, 0.41]	0.35^{***} [0.32, 0.38]	0.31^{***} [0.28, 0.34]	0.23^{***} [0.19, 0.26]
2.焦虑	55.47(9.27)	0.40^{***} [0.37, 0.43]	0.45^{***} [0.42, 0.47]	0.38^{***} [0.35, 0.41]	0.07^{***} [0.04, 0.11]
3.生活满意度	19.16(6.63)	-0.23^{***} [-0.27, 0.20]	-0.31^{***} [-0.34, -0.28]	-0.17^{***} [-0.21, -0.14]	-0.01 [-0.05, 0.02]
4.创造力倾向	173.70(18.56)	-0.03 [-0.07, 0.00]	-0.20^{***} [-0.23, -0.16]	0.07^{***} [0.04, 0.11]	0.14^{***} [0.11, 0.18]
5.睡眠：睡眠时长 [a]	–	-0.11^{***} [-0.15, -0.08]	-0.06^{***} [-0.09, -0.02]	-0.17^{***} [-0.21, -0.14]	-0.05^{**} [-0.08, -0.01]
6.睡眠：充足睡眠	3.37(0.83)	-0.16^{***} [-0.20, -0.13]	-0.15^{***} [-0.18, -0.11]	-0.21^{***} [-0.25, -0.18]	-0.03 [-0.06, 0.01]
7.睡眠：失眠问题 [a]	–	0.12^{***} [0.08, 0.15]	0.10^{***} [0.06, 0.13]	0.14^{***} [0.11, 0.18]	0.05^{*} [0.01, 0.08]

注：$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$

方括号内的数据代表该相关系数的 95% CI。

[a] 分类变量，采用斯皮尔曼相关。

3.4 最终版大学生学业内卷问卷

研究 2 基于项目分析、探索性因子分析和验证性因子分析结果，确定了最终版大学生学业内卷问卷。最终版问卷共 16 道题，包含三个子维度：舍心逐物、竞争失格和脱实务虚。进一步的内部一致性和结构效度分析结果显示，最终版问卷及三个子维度均具有良好的信效度。

4 研究 3：问卷的重测信度检验

4.1 研究目的

研究 3 对新的样本进行间隔一个月的重复测量，检验最终版问卷的重测信度。

4.2 研究方法

4.2.1 研究对象

112 名在读本科生参与了重复测量，其中 6 名被试未在规定的间隔一个月时间内填写问卷，7 名被试未按照要求完成问卷中添加的注意力检测题，其数据未被纳入分析。最终保留有效被试 99 人，其中男性 29 人(29.29%)，包含了大一(57 人)、大二(11 人)、大三(30 人)和大四(1 人)不同年级的学生。

4.2.2 研究工具

最终版大学生学业内卷问卷。

4.3 研究结果

最终版问卷间隔一个月的重测信度为 0.83 ($p < 0.001$, 95% CI = [0.76, 0.89])，三个子维度的重测信度分别为：0.79 ($p < 0.001$, 95% CI = [0.71, 0.86])，0.77 ($p < 0.001$, 95% CI = [0.68, 0.84])，0.80 ($p < 0.001$, 95% CI = [0.71, 0.86])，均高于 0.7，表明最终版问卷重测信度良好。

对前测、后测数据分别进行了内部一致性分析。在前测数据中，总问卷的 α 内部一致性系数为 0.84，三个子维度的 α 内部一致性系数分别为：0.77，0.71，0.71，均高于 0.7。在后测数据中，总问卷及三个子维度的 α 内部一致性系数分别为：0.89，0.85，0.74，0.78，均高于 0.7。结果表明，最终版问卷题目信度良好。

5 讨论

本研究结合文献分析析取的内卷定义，通过题目编制和信效度检验确定了 16 个条目、三因子的大学生学业内卷问卷。大样本的结果显示该问卷的校标关联效度合理：学业内卷总分较高的学生有着更强的零和信念、更高的特质焦虑水平、更低的生活满意度、更差的睡眠质量，但却没有显示出更高的创造力倾向。三个因子维度根据题目内容分别被命名为：舍心逐物——努力追求外在结果而丧失学习初心，竞争失格——因资源有限而过度竞争，以及脱实务虚——不研究如何求学问而一味追求高绩点。无独有偶，苑津山和幸泰杞 (2021)所提出的大学生内卷体现与本研究的三因子内容不谋而合，在内容层面佐证了学业内卷问卷的合理性。

理论方面，本研究深化了对内卷概念的理解。本文依据 Goldenweiser、Geertz 和黄宗智对内卷现象的描述将（1）既定框架的限制、（2）不断攀升的投入、（3）缺乏实质收益或发展作为内卷的核心特点。三个核心特点分别与内卷的不同描述相兼容。例如：刘世定和邱泽奇所言的“外部扩张受限”和计亚萍(2010)的“外部限制、自我锁定”都体现了既定框架的限制；范志海(2004)的“内部不断精细化”和王琛的“自我重复”都是“在既定框架内部投入攀升”的某

种形式；而夏骏的“无发展的增长”和赵祥辉的“有数量增长而无实质创新”也都体现了“缺乏实质收益和发展”这一特点。陈诚和包雷对内卷特点的总结也得到了三个类似的特点：（1）有限空间、（2）投入增加、（3）不能实现个人和总量上的共赢。从逻辑上而言，三个核心特点也缺一不可。“投入攀升却缺乏实质收益和发展”是人们内卷的直观感受，而“既定的框架限制”则是导致这种现象的先决条件。由此可见，将内卷定义为“一种在既定框架内部投入不断攀升却缺少实质收益或发展的现象”兼具合理性与实用性，是有益的定义参考。

应用方面，本研究开发的问卷为后续研究学业内卷提供了更为可靠的测量工具。以往的学业内卷问卷存在着内卷定义不明、内容效度欠佳或信效度证据缺失等诸多问题(Yan et al., 2022; Yang et al., 2023; Yi et al., 2022)。Yi 等人的内卷定义强调了“对有限资源的非理性竞争”[4]，但并未限定这种竞争无法带来实质收益或发展。例如，其题目“我在空闲时间学习因为我想获得更多的知识从而提高自己的竞争力”。如果学生通过获取知识来提高竞争力，那么个体可以有实质的知识收益，而整体也可以在良性竞争中不断提升水平。Yan 等人的定义则更多地强调“非理性竞争下的投入攀升”[5]。例如，“我为了跟上别人的学习节奏而焦急的学习”和“当我不做大多数人都会做的事情时，我会感到不安”。这些题目都不符合内卷的概念。Yang 等人则既没有明确内卷定义也没有提供问卷的信效度证据。其问卷中的许多题目与内卷的定义相去甚远。例如：“我会在社交媒体（如 B 站、微博、小红书等）关注一些知识分享类账号提升自己，以免被别人落下”和“我每天早起晚归去学习，以免被别人落下”。综上可知，虽然 Yi 等人和 Yang 等人都认为内卷是一种“非理性竞争”，但并未对“非理性”进行界定。徐英瑾(2021)曾批评内卷的指代太宽泛，认为不应该“泛指任何过度竞争的社会现象”。徐呼吁用严格的判断标准来限定内卷概念的边界，以免内卷一词因应用范围过于宽泛而失去语用价值。按此要求，本研究的学业内卷问卷不仅基于界定清晰的概念而且有着充足的信效度，也因此更有益于规范内卷研究。

现实意义方面，本研究显示大学生的学业内卷问题值得教育界和社会重视。全国范围的 3000 人数据显示，大学生学业内卷不仅存在而且相当普遍。若以总分高于 48 分（题目的平均评分高于 3 分，即多数评分为“符合”或“完全符合”）的个体为内卷者，那么 3000 人中有 1930 人符合标准，占比 64.33%。同样以题目平均分高于 3 分为标准，3000 人中舍心逐物者

[4] We define involution as types of behaviors wherein individuals are under peer pressure or extrinsic motivation to be involved in active or submissive irrational competition for limited social resources.

[5] We submit that academic involution behavior is more appropriately expressed as a behavior mode of increasing burden, excessive competition, and inefficient learning adopted by university students in the face of an uncertain and irrationally competitive environment.

占比 61.03%；竞争失格者占比 42.83%；脱实务虚者占比 71.33%。学业内卷的负面影响也十分明显。高内卷的个体焦虑水平更高、生活满意度更低、睡眠质量更差。内卷子维度舍心逐物更高的学生创造力倾向也更低，有悖于高等教育的初衷。虽然竞争失格有可能促进了学生的创造力倾向，但这种促进或许是以牺牲睡眠为代价的。由于睡眠不足会降低记忆力和免疫力，竞争失格的这种作用不仅难以持续，还可能会增加个人和社会的医疗成本(Gais et al., 2006; Prather et al., 2015)。高脱实务虚的学生虽然也有着更高的创造力倾向，但他们恐怕难以取得创造性的成就。创造性成就不仅依赖于个体创造力倾向，还依赖于顽强的决心和无止境的钻研(Zhang et al., 2020)，而后者可能是脱实务虚者所欠缺的。在学业中，脱实务虚者的高创造力可能体现在能够找到取得更好成绩的“捷径”。3000 人的数据显示，高脱实务虚者反而有着更靠前的成绩排名（$|\rho| = 0.14$, $p < 0.001$）。这印证了苑津山和幸泰杞“评价失真”的论断，也警示教育界当前的评价体系有一些问题值得反思。例如：我们的评价体系是否一定程度上“培养”了脱实务虚者？当脱实务虚的学生们步入工作后会给社会带来何种影响？

对于许多内卷的研究问题，本论文虽有推进，但远未完善。限于篇幅，本文就三个较重要问题进行补充说明。其一，判断是否内卷应该依赖于客观标准还是主观标准？社会学的研究者们更倾向于根据较为客观的边际效率来判断是否内卷。但是这一方式在现实中往往难以适用；尤其是针对许多依赖于个体主观感受的方面（例如：学习收益）。从易于应用的角度考虑，本研究选择基于个体的主观感受设计的学业内卷。其假设是：人们的主观体验能够较敏感地反映客观的系统状况。其二，内卷与非内卷的界限难以明确。黄宗智(2002)曾提出，内卷发生于劳动力边际产量开始递减之后。刘世定和邱泽奇对此进行了反驳，认为黄宗智的界定既不符合 Geertz 的定义，也并不现实。界定的困难在学业内卷中尤为明显；目前难以形成一个公认的内卷与非内卷的分界线。因此，本研究的问卷只区分出高低学业内卷的个体，而划定分界线的问题则留待后人。其三，内卷的反面尚难定论。本研究结合文献与研究者普遍认可的三个特点构建了内卷的定义。而内卷的反面到底是三个特点上均相反的现象还是只在其中部分特点上相反的现象？目前仍难以回答。本问卷初始版所设计的反向计分题尝试测量的是“没有明显封闭空间，投入攀升但有实质收益”的现象，但均没有发挥出符合预期的作用。内卷的反面应该是什么？“躺平”是否更适合作为内卷的反面？这些问题仍有待于未来研究进行探索。

6 结论

本研究开发的 16 题大学生学业内卷量表包含舍心逐物、竞争失格和脱实务虚三个维度。

3000 人的全国样本中，6 成以上的大学生认为自己符合内卷的描述。学业内卷量表得分较高的大学生有着更强的零和信念、更高的特质焦虑水平、更低的生活满意度、更差的睡眠质量，但却没有显示出更高的创造力倾向。可见，学业内卷在大学生教育中不仅普遍而且危害明显，值得社会与教育界的重视。

参 考 文 献

Chen, C., & Bao, L. (2022). The Origin of Involution and Solutions to Address Involution in Education. *Journal of China Examinations, 2*, 81-88.

[陈诚, 包雷. (2022).内卷的产生机制与教育内卷的破解. *中国考试, 2*, 81-88.]

Chen, H., Jackson, T., & Huang, X. (2006). The Negative Physical Self Scale: Initial development and validation in samples of Chinese adolescents and young adults. *Body Image, 3*(4), 401–412.

Crocker, J., & Canevello, A. (2008). Creating and undermining social support in communal relationships: The role of compassionate and self-image goals. *Journal of Personality and Social Psychology, 95*(3), 555–575.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71–75.

Eide, P. K., Vinje, V., Pripp, A. H., Mardal, K.-A., & Ringstad, G. (2021). Sleep deprivation impairs molecular clearance from the human brain. *Brain, 144*(3), 863–874.

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning & memory, 13*(3), 259–262.

Gao, M. (2014). Mediating Effect of Academic Emotion between College Adjustment and Learning Burnout. *Chinese Journal of Clinical Psychology, 22*(3), 537–539.

[高明. (2014). 学业情绪在大学适应与学习倦怠间的中介作用. *中国临床心理学杂志, 22*(3), 537–539.]

Geertz, C. (1963). *Agricultural involution: The processes of ecological change in Indonesia.* Univ of California Press.

Goodman, E., Adler, N. E., Kawachi, I., Frazier, A. L., Huang, B., & Colditz, G. A. (2001). Adolescents' perceptions of social status: development and evaluation of a new indicator. *Pediatrics, 108*(2), 1–8.

Huang, Z. Z. (2021). Re-thinking Involution and De-involution. *Open Times, 1*, 157–168.

[黄宗智. (2021). 再论内卷化, 兼论去内卷化. *开放时代, 1*, 157–168.]

Jeffreys, H. (1961). *The theory of probability (3rd ed).* New York: Oxford university press.

Liu, Y., Wu, B., Petti, C., Wu, X., & Han, S. (2018). Self-construals moderate associations between trait creativity and social brain network. *Neuropsychologia*, *111*, 284-291.

McCullough, C. (2019). Review of "agricultural involution: the processes of ecological change in Indonesia" by Clifford Geertz. *International Journal of Anthropology and Ethnology, 3*, 1–5.

Olatunji, B. O., Williams, N. L., Tolin, D. F., Abramowitz, J. S., Sawchuk, C. N., Lohr, J. M., & Elwood, L. S. (2007). The Disgust Scale: Item analysis, factor structure, and suggestions for refinement. *Psychological Assessment, 19*(3), 281–297.

Pang, S. X., & Li, S. X. (2010). Involution：A Description of Bottleneck in Education. *Higher Education Development and Evaluation, 26*(6), 24–29.

[庞守兴, 李书肖. (2010). 内卷化:教育的瓶颈描述. *高教发展与评估, 26*(6), 24–29.]

Prather, A. A., Janicki-Deverts, D., Hall, M. H., & Cohen, S. (2015). Behaviorally assessed sleep and susceptibility to the common cold. *Sleep, 38*(9), 1353–1359.

Vagos, P., Rijo, D., & Santos, I. M. (2016). Scenes for Social Information Processing in Adolescence: Item and factor analytic procedures for psychometric appraisal. *Psychological Assessment, 28*(4), 416–428.

Wang, C. (2013). Involution and its Cultural Psychological Mechanism. *Journal of Shenzhen University (Humanities & Social Sciences), 5*, 182–190.

[王琛. (2013). "内卷化"及其文化心理机制分析. *深圳大学学报 (人文社会科学版), 5*, 182–190.]

Xu, Y. J. (2021). Digital Fetishism: The Essence of Involution. *Exploration and Free Views, 3*, 57–65

[徐英瑾. (2021). 数字拜物教:"内卷化"的本质. *探索与争鸣, 3*, 57–65.]

Yan, D., Zhang, H., Guo, S., & Zeng, W. (2022). Influence of anxiety on university students' academic involution behavior during COVID-19 pandemic: Mediating effect of cognitive closure needs. *Frontiers in psychology,* 13.

Yang, Y., Peng, Y., Li, W., Lu, S., Wang, C., Chen, S., & Zhong, J. (2023). Psychometric evaluation of the academic involution scale for college students in China: An application of Rasch analysis. *Frontiers in psychology,* 14.

Yi, D., Wu, J., Zhang, M., Zeng, Q., Wang, J., Liang, J., & Cai, Y. (2022). Does Involution Cause Anxiety? An Empirical Study from Chinese Universities. *International Journal of Environmental Research and Public Health, 19*(16), 9826.

Zhao, T. T. & Hu, Y. T. (2022). The Ability and Response: The Function of the Grade Point Average System from the Undergraduate Transcripts. *University Education Science, 3*, 22–30.

[赵婷婷, 胡艳婷. (2022). 能为与应为:从本科生成绩单看绩点制的功能. *大学教育科学, 3*, 22–30.]

Zhao, X. H. (2021). "Involution" of the Doctoral Publication System: The Characterization, Mechanism and Correction. *Journal of Higher Education Management, 15*(3), 104–113.

[赵祥辉. (2021). 博士生发表制度的 “内卷化”: 表征, 机理与矫治. *高校教育管理, 15*(3), 104–113.]

Zhang, Q. & Gao, Y. Y. (2022). Competition and Game: The Academic Returns and Mental Health Costs of Private Tutoring. *Chinese Journal of Sociology, 42*(3), 159–194.

[张骞, 高雅仪. (2022). 竞争与博弈:课外补习的学业回报与心理健康代价. *社会, 42*(3), 159–194.]

Zhang, Z. S., Hoxha, L., Aljughaiman, A., Arnliu, A., & Ziegler, A. (2020). Social Environmental Factors and Personal Motivational Factors Associated with Creative Achievement: A Cross-cultural Perspective. *The Journal of creative behavior* (1).

陈雪飞. (2019). 应试教育背景下大学生学习心理分析. *校园心理, 17*(3), 210–212.

戴子涵. (2022). 培养游戏精神: 突破教育内卷重围之道. *少年儿童研究, 1*, 66–72.

范志海. (2004). 论中国制度创新中的“内卷化”问题. *社会, 4*, 4–7.

郭继强. (2007). “内卷化”概念新理解. *社会学研究, 3*, 194–208.

黄宗智. (2002). 发展还是内卷? 十八世纪英国与中国——评彭慕兰《 大分岔: 欧洲, 中国及现代世界经济的发展》. *历史研究, 4*, 149–176.

计亚萍. (2010). “内卷化”理论研究综述. *长春工业大学学报 (社会科学版), 22*(3), 48–49.

刘世定, 邱泽奇. (2004). “内卷化”概念辨析. *社会学研究, 5*, 96–110.

汪卫东, 王希林, 马弘. (1999). *心理卫生评定量表手册 (增订版 pp. 235-238)*. 北京：中国心理卫生杂志社.

夏骏. (2012). 高等教育内卷化增长探讨. *教育评论, 3*, 12–14.

熊丙奇. (2021). 实行十年义务教育可缓解"内卷"? *在线学习, 8*, 79.

苑津山, 幸泰杞. (2021). 入局与破局”: 高校学生内卷参与者的行为逻辑与身心自救. *高教探索, 10*, 123–128.